

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-042890

(43)Date of publication of application : 16.02.2001

(51)Int.Cl.

G10L 15/22
G06F 3/16

(21)Application number : 11-217073

(71)Applicant : TOSHIBA TEC CORP

(22)Date of filing : 30.07.1999

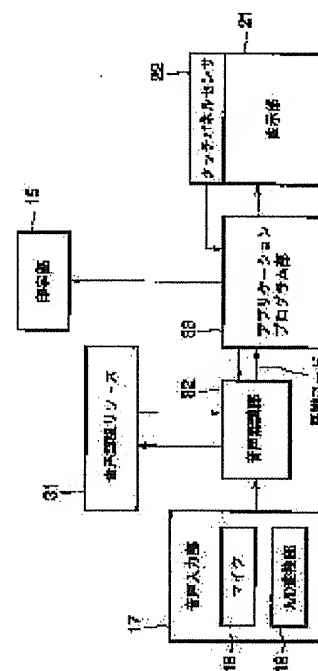
(72)Inventor : TAKAHASHI TAKAHIDE
YAMAMOTO KENICHI

(54) VOICE RECOGNIZING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a voice recognizing device which is capable of simply selecting an arbitrary input column, performing voice input and preventing unnecessary voice input, and is excellent in use convenience.

SOLUTION: This voice recognizing device is provided with a voice input part 17 for inputting the voice of speakers, a voice recognition resource 31 which stores words and phrases to be recognized beforehand, a voice recognizing part 32 which recognizes the words and phrases which are inputted by the voice input by extracting the words and phrases from among the same of the voice recognizing resource when the voice is inputted at an input state of voice, a display part 21 which displays buttons which are respectively related to the plural data input columns and each input column, and a touch panel sensor 22 which is overlapped and disposed on the screen of the display part 21 and detects the push down states of the respective buttons displayed on the display part 21. Therein, the device is set to be the input state of voice in accordance with the push down state of the respective buttons detected by the touch panel sensor 22, the results which are recognized on the voice recognition part are displayed in the data input columns which are related to the push down buttons and, at the same time, are inputted as data of the data input columns.



(43)公開日 平成13年2月16日(2001.2.16)

(51)Int.Cl. ⁷	識別記号	F I	テマコード*(参考)
G 1 0 L 15/22		G 1 0 L 3/00	5 6 1 C 5 D 0 1 5
G 0 6 F 3/16	3 2 0	G 0 6 F 3/16	3 2 0 H 9 A 0 0 1
			3 2 0 A
			3 2 0 D
		G 1 0 L 3/00	5 6 1 E
審査請求 未請求 請求項の数6 O L (全 7 頁)			

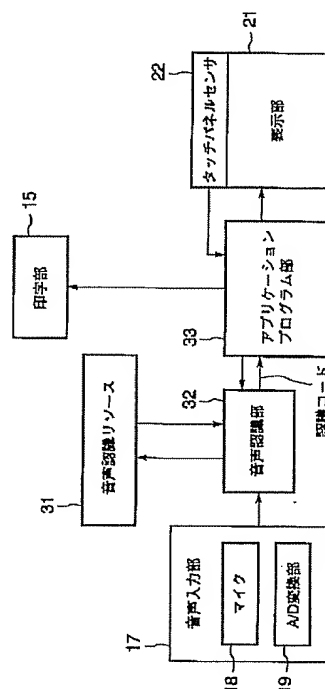
(21)出願番号	特願平11-217073	(71)出願人	000003562 東芝テック株式会社 東京都千代田区神田錦町1丁目1番地
(22)出願日	平成11年7月30日(1999.7.30)	(72)発明者	高橋 隆英 静岡県三島市南町6番78号 東芝テック株 式会社三島事業所内
		(72)発明者	山本 健一 静岡県三島市南町6番78号 東芝テック株 式会社三島事業所内
		(74)代理人	100058479 弁理士 鈴江 武彦 (外6名) Fターム(参考) 5D015 AA02 BB01 LL05 LL08 9A001 BB03 EE02 EE04 EE05 FF03 HH17 JJ35 KK31 KK42

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【課題】 任意の入力欄を簡単に選択して音声入力することができ、余計な音声入力を防止できる使い勝手のよい音声認識装置を提供する。

【解決手段】 話者の音声を入力するための音声入力部 1 7 と、予め認識されるべき語句を記憶した音声認識リソース 3 1 と、音声入力状態のときに音声を入力すると、音声認識リソースの語句の中から抽出することにより、音声入力した語句を認識する音声認識部 3 2 と、複数のデータ入力欄と各データ入力欄にそれぞれ関連づけられたボタンを表示する表示部 2 1 と、表示部の画面上に重ねて設けられ、その表示部に表示した各ボタンの押下状態を検出するタッチパネルセンサ 2 2 と、タッチパネルセンサが検出した各ボタンの押下状態に応じて音声入力状態にし、音声認識部で認識された結果を押下されたボタンに関連づけられたデータ入力欄へ表示するとともにそのデータ入力欄のデータとして入力する。



【特許請求の範囲】

【請求項 1】 話者の音声を入力するための音声入力手段と、
予め認識されるべき語句を記憶した音声認識リソースと、
音声入力状態のときに前記音声入力手段から音声を入力すると、前記音声認識リソースの語句の中から抽出することにより、音声入力した語句を認識する音声認識手段と、
複数のデータ入力欄と各データ入力欄にそれぞれ関連づけられたボタンを表示する表示手段と、
前記表示手段の画面上に重ねて設けられ、その表示手段に表示した各ボタンの押下状態を検出するタッチパネルセンサと、
前記タッチパネルセンサが検出した各ボタンの押下状態に応じて前記音声認識手段を音声入力状態にし、この音声認識手段で認識された結果を押下されたボタンに関連づけられたデータ入力欄へ表示するとともにそのデータ入力欄のデータとして入力する音声入力制御手段と、
を設けたことを特徴とする音声認識装置。
【請求項 2】 前記音声入力制御手段は、前記タッチパネルセンサにより前記ボタンが押されたと判断している間は、前記音声認識手段を音声入力状態にし、前記ボタンが離されたと判断したときに音声入力状態を終了することを特徴とする請求項 1 記載の音声認識装置。
【請求項 3】 前記音声入力制御手段は、前記タッチパネルで検出された各ボタンの押下状態に基づいて、前記ボタンが一度押されたと判断したときは音声入力状態にし、もう一度押されたと判断した場合は音声入力状態を終了することを特徴とする請求項 1 記載の音声認識装置。
【請求項 4】 話者の音声を入力するための音声入力手段と、
予め認識されるべき語句を記憶した音声認識リソースと、
音声入力状態のときに前記音声入力手段から音声を入力すると、前記音声認識リソースの語句の中から抽出することにより、音声入力した語句を認識する音声認識手段と、
複数のデータ入力欄を表示する表示手段と、
この表示手段の各データ入力欄にそれぞれ関連づけられたボタンと、
前記各ボタンの押下状態を検出するボタン状態検出手段と、
前記ボタン状態検出手段が検出した各ボタンの押下状態に応じて前記音声認識手段を音声入力状態にし、この音声認識手段で認識された結果を押下されたボタンに関連づけられたデータ入力欄へ表示するとともにそのデータ入力欄のデータとして入力する音声入力制御手段と、
を設けたことを特徴とする音声認識装置。

【請求項 5】 前記音声入力制御手段は、前記ボタン状態検出手段により前記ボタンが押されたと判断している間は、前記音声認識手段を音声入力状態にし、前記ボタンが離されたと判断したときに音声入力状態を終了することを特徴とする請求項 4 記載の音声認識装置。

【請求項 6】 前記音声入力制御手段は、前記ボタン状態検出手段で検出された各ボタンの押下状態に基づいて、前記ボタンが一度押されたと判断したときは音声入力状態にし、もう一度押されたと判断した場合は音声入力状態を終了することを特徴とする請求項 4 記載の音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、表示画面に設けた入力欄に音声でデータ入力を行う音声認識装置に関する。

【0002】

【従来の技術】従来の音声認識装置は、図 7 に示すように音声を入力するマイク 1 a とこのマイクからの音声を変換する A/D 変換器 1 b を備える音声入力部 1、予め認識されるべき語句と各語句に対して定義した認識コードからなる音声認識リソース 2、この音声入力部 1 からの出力に基づいて語句を認識し、その語句に対応する認識コードを音声認識リソース 2 に基づいて抽出する音声認識部 3、複数の入力欄を表示させる表示部 4、音声入力により入力する入力欄を選択するキー操作など操作者が各種のキー操作を行うためのキーボード 5、ポインタデバイスとしてのマウス 6、キーボード 5 やマウス 6 により入力欄が選択されると音声認識部 3 を音声入力可能状態にする命令を出力し、音声入力により音声認識部 3 からの認識コードに基づいて商品名や金額の入力を行うアプリケーションプログラム部 7 から構成される。

【0003】上記表示部 4 は、音声入力を行う場合には図 8 に示すような画面表示を行うようになっている。この表示画面には複数の入力欄がある。これらの入力欄の横にあるデータ 1、データ 2、…は、各入力欄に入力するデータの名称を示している。

【0004】このような装置において、音声入力を行う場合には、まず、表示部 4 に図 8 に示すような表示画面が表示される。そして、キーボード 5 やマウス 6 により入力欄が選択され、音声入力部 1 から音声が入力されると、音声認識部 3 で音声認識がなされ、認識コードが出力される。すると、アプリケーションプログラム部 7 は、音声認識部 3 から出力された認識コードに基づいて得られたデータを上記キーボード 5 やマウス 6 により選択された入力欄のデータとして入力し、その結果を選択された入力欄に表示する。

【0005】また、入力欄を選択する際、上記キーボード 5 やマウス 6 を使用しなくても、最初はデフォルト値

としてデータ 1 の入力欄が選択されるようにしておき、特定の音声キーワードによって入力欄を選択するものもある。このような装置では、例えば句読点を示す「まる」、「ここで改行」、「次の欄移動」等の音声キーワードが入力されると次の入力欄に移り、そこに入力したいデータを発声すると当該入力欄にデータが入力されるようになっている。

【0006】

【発明が解決しようとする課題】ところで、音声認識装置の構成を必要最小限にして装置の小型化、コスト低下などを図るため、マウスやキーボードを設けないことがある。このような装置では、上述したような複数の入力欄を持たせる場合、マウスやキーボードを使って入力欄を選択することができないため、いったん音声入力されたデータが入力欄 1～n のうち、どれに該当しているのか装置側からは判別できないという問題がある。

【0007】また、音声キーワードで入力欄を特定させる場合、操作者はその入力欄の順番を意識して音声入力を行わなければならないなど操作者への負担が大きく、操作ミスの原因になるという問題がある。

【0008】そこで、本発明は、入力欄に関連づけられたボタンを表示し、タッチパネルセンサが検出したボタンの押下状態に応じて音声入力状態にすることによって、簡単に入力欄を選択することができる使い勝手のよい音声認識装置を提供しようとするものである。

【0009】

【課題を解決するための手段】請求項 1 の本発明は、話者の音声を入力するための音声入力手段と、予め認識されるべき語句を記憶した音声認識リソースと、音声入力状態のときに音声入力手段から音声を入力すると、音声認識リソースの語句の中から抽出することにより、音声入力した語句を認識する音声認識手段と、複数のデータ入力欄と各データ入力欄にそれぞれ関連づけられたボタンを表示する表示手段と、表示手段の画面上に重ねて設けられ、その表示手段に表示した各ボタンの押下状態を検出するタッチパネルセンサと、タッチパネルセンサが検出した各ボタンの押下状態に応じて音声認識手段を音声入力状態にし、この音声認識手段で認識された結果を押下されたボタンに関連づけられたデータ入力欄へ表示するとともにそのデータ入力欄のデータとして入力する音声入力制御手段とを設けたことを特徴とする音声認識装置である。

【0010】請求項 2 の本発明は、音声入力制御手段は、タッチパネルセンサによりボタンが押されたと判断している間は、音声認識手段を音声入力状態にし、ボタンが離されたと判断したときに音声入力状態を終了することを特徴とする請求項 1 記載の音声認識装置である。

【0011】請求項 3 の本発明は、音声入力制御手段は、タッチパネルで検出された各ボタンの押下状態に基づいて、ボタンが一度押されたと判断したときは音声入

力状態にし、もう一度押されたと判断した場合は音声入力状態を終了することを特徴とする請求項 1 記載の音声認識装置である。

【0012】請求項 4 の本発明は、話者の音声を入力するための音声入力手段と、予め認識されるべき語句を記憶した音声認識リソースと、音声入力状態のときに音声入力手段から音声を入力すると、音声認識リソースの語句の中から抽出することにより、音声入力した語句を認識する音声認識手段と、複数のデータ入力欄を表示する表示手段とこの表示手段の各データ入力欄にそれぞれ関連づけられたボタンと、各ボタンの押下状態を検出するボタン状態検出手段と、ボタン状態検出手段が検出した各ボタンの押下状態に応じて音声認識手段を音声入力状態にし、この音声認識手段で認識された結果を押下されたボタンに関連づけられたデータ入力欄へ表示するとともにそのデータ入力欄のデータとして入力する音声入力制御手段とを設けたことを特徴とする音声認識装置である。

【0013】請求項 5 の本発明は、音声入力制御手段は、ボタン状態検出手段によりボタンが押されたと判断している間は、音声認識手段を音声入力状態にし、ボタンが離されたと判断したときに音声入力状態を終了することを特徴とする請求項 4 記載の音声認識装置である。

【0014】請求項 6 の本発明は、音声入力制御手段は、ボタン状態検出手段で検出された各ボタンの押下状態に基づいて、ボタンが一度押されたと判断したときは音声入力状態にし、もう一度押されたと判断した場合は音声入力状態を終了することを特徴とする請求項 4 記載の音声認識装置である。

【0015】

【発明の実施の形態】以下、本発明の実施の形態を図 1 ないし図 6 を参照して説明する。図 1 は、本実施の形態に係る音声認識装置の構成を示すブロック図で、11 は制御部本体を構成する CPU（中央処理装置）、12 はこの CPU 11 が実行するプログラムデータを格納した ROM（リード・オンリ・メモリ）、13 は各種データ処理のために使用されるメモリ等を設けた RAM（ランダム・アクセス・メモリ）、14 はハードディスク装置（HDD）、15 は所定情報を印字してラベルの発行などを行う印字部、17 は音声をアナログ信号として入力するマイク 18 とこのマイク 18 からの音声をアナログ信号として入力した音声をデジタル信号に変換する A/D 変換器 19 を備えた音声入力手段としての音声入力部、20 は入力した音声を認識した結果やタッチパネルのボタンを表示する表示手段としての表示部 21 及びタッチパネルセンサ 22 を設けたタッチパネル付ディスプレイである。このタッチパネル付ディスプレイ 20 の表示部 21 は表示制御部 23 に接続しており、タッチパネルセンサ 22 はタッチパネルセンサ制御部 24 に接続している。

【0016】上記CPU11と、ROM12、RAM13、ハードディスク装置14、印字部15、A/D変換器19、表示制御部23、センサ制御部24とは、それぞれデータバス、制御バス、アドレスバスなどのバスラインで接続されている。

【0017】図2は、本実施の形態にかかる音声認識装置の構成を示す機能ブロック図であり、31は認識されるべき語句（数を含む）と各語句に対して定義された（関連づけられた）認識コードからなる音声認識リソース、32は音声入力部17からの出力に基づいて、入力した音声に対応する語句を認識し（音声認識手段）、その語句に対応する認識コードを音声認識リソース31から抽出して出力する音声認識部、33は音声認識部32からの認識コードに基づいて表示部21の入力欄（データ入力欄）に表示を行うとともにその入力欄のデータとして入力し、そのデータに基づいて商品名、商品の単価の登録などを行い、印字部15によりラベルの発行などを行うアプリケーションプログラム部である。上記音声認識リソース31は、各入力欄に入力するデータの種類のごとに設けられ、それぞれ各入力欄に関連づけられている。各音声認識リソース31には、該当する種類のデータについての予め認識されるべき語句とその語句に関係づけられた認識コードがそれぞれ記憶されている。

【0018】上記音声認識部32は、音声入力状態にあるときのみ、音声入力部17のマイク18から入力した音声認識して、認識コードをアプリケーションプログラム部33へ出力する。従って、上記音声認識部32は、音声入力状態にないときは、たとえ音声入力部17のマイク18から音声が入力されても、それを無視する。

【0019】また、音声認識部32は、アプリケーションプログラム部33から許可指令を受けたときに音声入力状態となり、終了指令を受けたときに音声入力状態を終了する。

【0020】なお、上記音声認識部32、アプリケーションプログラム部33は、具体的には例えばハードディスク装置14、ROM12などに記憶され、上記CPU11が読取可能なソフトウェアプログラムで構成される。

【0021】上記アプリケーションプログラム部33は、音声により商品名、単価などのデータを入力する場合には、表示部21に図5に示すような表示画面41を表示する。具体的には、複数の入力欄42、各入力欄42に関連づけられたボタン43、各入力欄42に入力するデータ名（データ1、データ2…）を各入力欄42に並べて表示する。

【0022】ここで、音声により商品名、単価などの入力を行う場合にアプリケーションプログラム部33においてCPU11が行う処理を図3に示すフローチャートに基づいて説明する。上記アプリケーションプログラム

部33では、タッチパネルセンサ22の出力により表示部21に表示したボタン43の押下状態を検出する（ボタン状態検出手段）する。例えば、ボタン43の状態フラグを設け、ボタン43が押下されたときには状態フラグを1とし、押されている間は、状態フラグを1に保持する。そして、ボタン43が離されたときは状態フラグを0とする。

【0023】そして、上記アプリケーションプログラム部33では、ボタン43の押下状態を監視しながら図3に示す入力処理を行う。まず、ST（ステップ）1にて状態フラグなどに基づいてボタン43が押されたかを判断する。ボタン43が押されたと判断した場合は、ST2にて押されたボタン43に関連づけられた入力欄42を選択する。

【0024】続いて、ST3にて当該入力欄42に関連づけられた音声認識リソース31に切替え、ST4にて音声認識部32に許可指令を行い、音声入力状態にする。この状態で、音声入力部17のマイク18から音声を入力すると、音声認識部32は、その音声に基づいてボタン43により選択された入力欄42の音声認識リソース31に基づいて音声認識を行い、認識コードをアプリケーションプログラム部33に出力する。

【0025】アプリケーションプログラム部33では、ST5にて音声認識部32から認識コードを受取ると、その認識コードにより得られたデータを当該入力欄42のデータとして入力し、当該入力欄42にそのデータを表示して（音声入力制御手段）、一連の入力処理を終了する。そして、入力処理がすべて終了するまで、この入力処理が繰返して実行される。なお、入力処理がすべて終了すると、アプリケーションプログラム部33は、その入力したデータに基づいて業務処理を行う。例えば、商品名、商品単価の登録などを行って、そのデータを印字データとして印字部15に送信する。これにより、印字部15は印字データに基づいて印字処理を行い、ラベルの発行等を行う。

【0026】上記アプリケーションプログラム部33においては、上記入力処理を行っている間に、状態フラグなどによりボタン状態を検出し（ボタン状態検出手段）、その結果に基づいてボタン43が離されたと判断した場合は、図4に示すような割込処理を行う。この割込処理では、音声認識部32に終了指令を行い、音声入力状態を終了する。これにより、音声認識部32は、音声入力状態を終了した後に音声入力部17から音声が入力されても、それを無視する。

【0027】なお、本実施の形態においては、各入力欄42に入力するデータ名を各入力欄42に並べて表示する場合について述べたが、図6に示すように各ボタン43上にデータ名を表示してもよい。

【0028】このような構成の本発明の実施の形態においては、例えばラベルに印刷する商品名を各種類ごとに

音声入力する場合、表示部に図 6 に示すような画面が表示される。各入力欄 4 2 に並べてボタン 4 3 を配置し、各ボタン 4 3 上には各入力欄 4 2 に入力する商品名（魚類、野菜類、肉類…）を表示してある。

【0029】例えば、魚類の商品名の入力欄 4 2 に音声入力を行う場合は、魚類のボタン 4 3 を押すと、魚類の音声認識リソースが音声認識リソースが選択されて音声入力状態になる。そして、その魚類のボタン 4 3 を押しながら、マイク 1 8 に向けて「ぶり」と発声すると、音声認識されて、魚類の入力欄 4 2 のデータとして「ぶり」が入力され、入力欄 4 2 に「ぶり」が表示される。その後、ボタン 4 3 を離すと、音声入力状態が終了し、印字部 1 5 により「ぶり」と印字されたラベルが発行される。

【0030】このように、表示部に各入力欄 4 2 とこの入力欄 4 2 に関連づけられたボタン 4 3 を表示し、タッチパネルセンサ 2 2 でそのボタン 4 3 の押下状態を監視し、ボタン 4 3 が押下している間は音声入力状態にしてマイク 1 8 から入力した音声の認識を行ってその結果をその入力欄 4 2 のデータとして入力するとともに、その入力欄 4 2 に表示し、ボタン 4 3 を離したときは音声入力状態を終了することにより、キーボードやマウスがなくても、簡単に入力欄 4 2 を選択して音声入力することができるとともに、ボタン 4 3 を押している間だけ音声入力状態にするので、操作者側で発声のタイミングをとることが容易となる使い勝手のよい音声認識装置を提供できる。

【0031】また、ボタン 4 3 を押している間だけ音声入力状態にするので、不要な音声は認識されことなく、必要な音声のみについて認識を行うことができるため、認識率が向上する。また、複数の入力欄 4 2 があっても、任意の入力欄 4 2 にデータを入力することができる。これにより、操作者側で入力欄 4 2 の順番を意識して入力を行う必要がなくなるので操作者側の負担を軽減することができる。

【0032】また、ボタン 4 3 の操作により入力欄 4 2 ごとに関連づけられた音声認識リソースを切替えることができるので、認識率が向上するとともに、音声認識部の処理量を軽減でき、検出時間を短縮できる。

【0033】なお、本実施の形態では、ボタン 4 3 を押すと音声入力状態になり、離すと音声入力状態が終了するようにしたが、必ずしもこれに限定されるものではなく、ボタン 4 3 を 1 回押すと音声入力状態になり、そのボタン 4 3 をもう一度押すと音声入力状態が終了するようにしてもよい。

【0034】また、ボタン 4 3 は必ずしも表示部 2 1 の表示画面上に表示される必要はなく、各ボタン 4 3 を表示部 2 1 の表示画面の近傍に別途設けたり、キーボードを有する装置においては各ボタン 4 3 をキーボード上に割り当てて、各ボタン 4 3 の押下状態を状態フラグなど

で検出し（ボタン状態検出手段）、各ボタンの押下状態に応じて音声認識部 3 2 を音声入力状態にしてもよい。このようにしても同様の効果を得られる。

【0035】

【発明の効果】以上詳述したように本発明によれば、表示部に各入力欄とこの入力欄に関連づけられたボタンを表示し、タッチパネルセンサ又はボタン状態検出手段によるボタンの押下状態に応じてボタンが押されている間は音声入力状態にして入力した音声の認識を行ってその結果をその入力欄のデータとして入力するとともに、その入力欄に表示することにより、キーボードやマウスがなくても簡単に入力欄を選択して音声入力することができる。

【0036】また、ボタンを押している間だけ音声入力状態にし、ボタンを離すと音声入力状態を終了するので、操作者側で発声のタイミングをとることが容易となり、さらに不要な音声は認識されことなく、必要な音声のみについて認識を行うことができるため、認識率が向上する。また、音声入力制御手段として一度ボタンを押すと音声入力状態にして、もう一度ボタンを押すと音声入力状態を終了するようにしても、同様の効果が得られる。

【0037】また、複数の入力欄があっても、任意の入力欄にデータを入力することができる。これにより、操作者側で入力欄の順番を意識して入力を行う必要がなくなるので操作者側の負担を軽減することができる。

【図面の簡単な説明】

【図 1】本発明の実施の形態に係る音声認識装置の構成を示すブロック図。

【図 2】本実施の形態における機能ブロック図。

【図 3】本実施の形態における入力処理を示す流れ図。

【図 4】本実施の形態における割込処理を示す流れ図。

【図 5】本実施の形態における表示部の表示例を示す流れ図。

【図 6】本実施の形態における表示部の他の表示例を示す流れ図。

【図 7】従来の音声認識装置の機能ブロック図。

【図 8】従来の音声認識装置における表示部の表示例を示す流れ図。

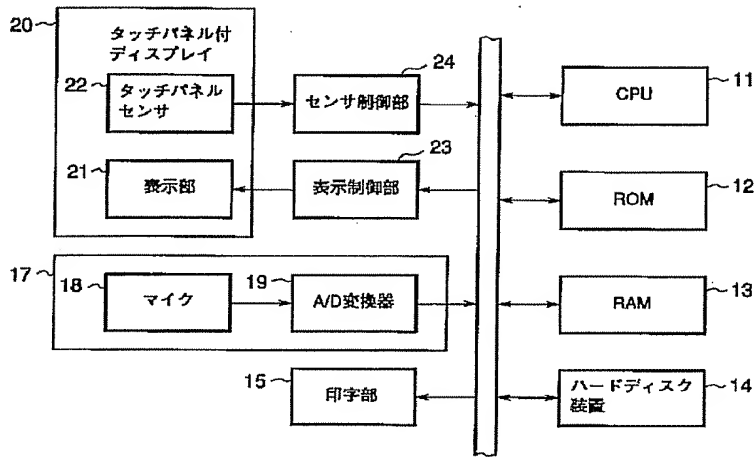
【符号の説明】

- 1 1 … CPU
- 1 2 … ROM
- 1 3 … RAM
- 1 7 … 音声入力部
- 1 8 … マイク
- 2 1 … 表示部
- 2 2 … タッチパネルセンサ
- 3 1 … 音声認識リソース
- 3 2 … 音声認識部
- 3 3 … アプリケーションプログラム部

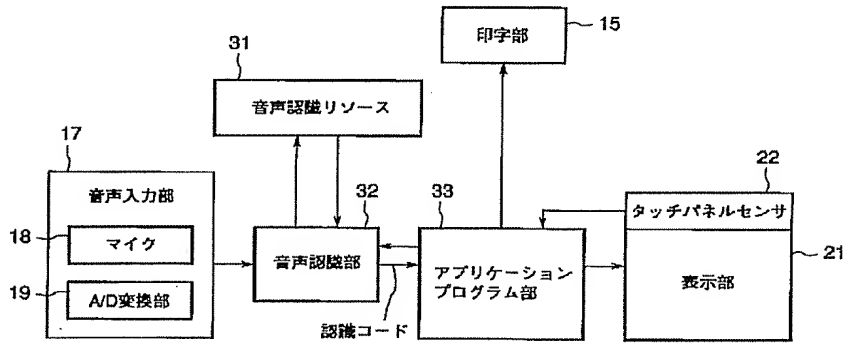
4 2 …入力欄

4 3 …ボタン

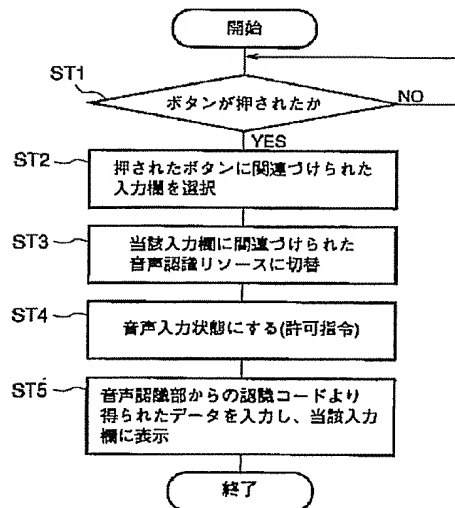
【図 1】



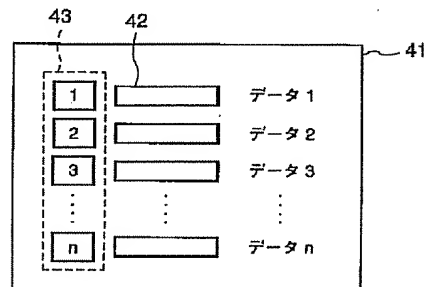
【図 2】



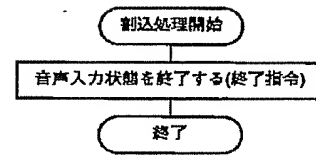
【図 3】



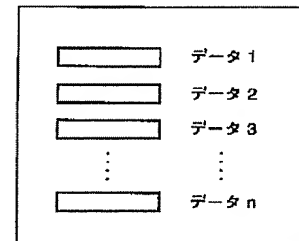
【図 5】



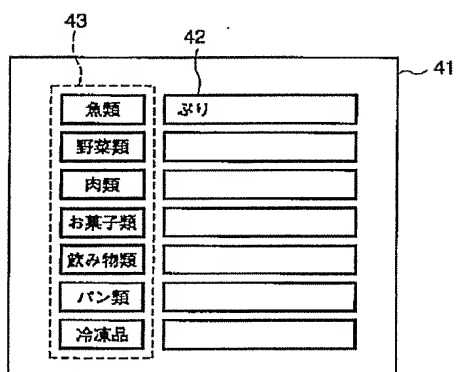
【図 4】



【図 8】



【図 6】



【図 7】

